

Supplement to A tug-of-war between driver and passenger mutations in cancer and other adaptive processes

Christopher D McFarland*, Leonid Mirny*[†], and Kirill S Korolev[‡]

*Graduate Program in Biophysics, Harvard University, Boston, MA 02115

[†]Department of Physics & Institute for Medical Engineering and Science, Massachusetts Institute of Technology (MIT), Cambridge, MA 02139, USA

[‡]Department of Physics and Program in Bioinformatics, Boston University, Boston, MA 02215

September 24, 2014

Contents

1	Mathematical description and analysis of our model of advantageous drivers and deleterious passengers.	3
1.1	Detailed formulation of our model	3
1.2	Parameter range investigated	6
1.3	A two-hit model of cancer progression	8
2	An analytical model of dynamics	9
2.1	Estimating the probability of cancer	10
2.2	Estimating the mean time to progression	12
2.3	Accumulation of deleterious passengers	14
2.4	Effects of deleterious passengers on fixation probability and clone fitness of drivers	16
3	A traditional model of cancer progression with drivers and neutral passengers.	18
4	Materials & Methods	22
4.1	Simulations	22
4.2	Inference of lesion formation rate	22
4.3	Analysis of cancer genomes	22

List of Figures

S1	Age-incidence curves in our model depend primarily on s_d and match observed age-incidence curves well at mid/late life	29
S2	Somatic Nonsynonymous Mutations (SNMs) and Somatic Copy Number Alterations (SCNAs) exhibit similar positive linear relationships among cancer subtypes	30
S3	Distribution of mutation totals in cancers is highly dispersed and positively skewed	31
S4	Analytical framework predicts probability of cancer across parameter space	33
S5	Simulations exhibit path independence	34
S6	Combination treatments that increase mutation rate and selection against passengers work best	35
S7	A two-hit driver model with deleterious passengers also experiences a population size-dependent barrier to progression	36

List of Tables

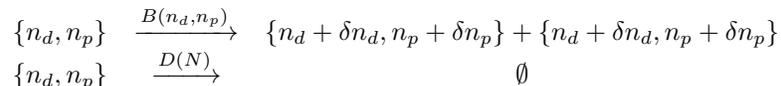
S1	Evolutionary parameters explored in this study	37
S2	Average number of driver and passenger mutations by tumor type	38
S3	Linear relationship between drivers and passengers cannot be explained by other tumor properties	39
S4	Kolmogorov-Smirnov goodness of fit estimates of s_d (our model) and k (traditional model) for sequenced cancer subtypes	41

1 Mathematical description and analysis of our model of advantageous drivers and deleterious passengers.

In this section, we present an exact mathematical formulation of our model, briefly explore a ‘two-hit’ model where the first driver mutation confers no fitness advantage, describe the broad ranges of parameters that we chose to explore, and offer an analytical description of our model. In our analytical analysis, we estimate three properties of the model that were not completely addressed in the main text: (i) the effects of stochasticity on population dynamics, (ii) the rate of accumulation of deleterious passengers, and (iii) the interference of driver accumulation by deleterious passengers.

1.1 Detailed formulation of our model

As mentioned in the main text, we model cancer progression via a first-order Gillespie Algorithm. Each cell within the cancer is represented by a separate “chemical species” or reactant in the Gillespie algorithm. Cells are defined by their state: $\{n_d, n_p\}$ ¹. n_d denotes the number of drivers in the cell, while n_p denotes the number of passengers. Cells can then divide, with and without mutations, and die according to the following half-reactions:



The functions $B(n_d, n_p)$ and $D(N)$ represent the birth and death rates of cells, while N represents the total number of cells in the precancerous population. The birth rate assumes multiplicative fitness effects of mutations and no epistasis between mutations:

$$B(n_d, n_p) = \frac{(1 + s_d)^{n_d}}{(1 + s_p)^{n_p}} \approx (1 + s_d)^{n_d} (1 - s_p)^{n_p} \quad [\text{S1}]$$

We also define a generation in terms of the mean division time:

$$1 \text{ generation} = \frac{1}{1/N \sum_{i=0}^N B(n_{d_i}, n_{p_i})}$$

¹Depending upon the number of cells and genomes in the population, it may be more efficient to model cancer as a set of genotypes that gain and lose cells (rather than a set of cells that gain mutations, as we have done). The efficiency of this choice depends upon whether the number of genomes in the population outnumber the number of cells, or vice versa. However, the efficiency of a Gillespie algorithm depends very weakly (logarithmically) on the number of chemical species [1]. More importantly, however, the other steps in the simulation (creating mutations, and calculating birth/death rates) are faster for individual cells than for genomes. Thus, using cells as the basic element of simulations, rather than genomes, is faster even under circumstances where the number of cells greatly outnumber the number of genomes. This design choice also allows more plasticity in model design and allows tracking of coalescent trees.

The death rate is defined such that, in the absence of mutations, the expectation value of the population size will obey a Gompertz curve at large sizes and a logistic curve at small sizes:

$$D(N) = \text{Log}\left[1 + \frac{(e-1)N}{N^0}\right]$$

We used a simpler form of this death function for populations grown to less than 10^6 cells:

$$D(N) = \frac{N}{N^0} \quad [\text{S2}]$$

This second functional form was used in much of our analysis because (i) it did not significantly alter dynamics at small sizes [2], (ii) it has been used previously [3], (iii) it is easier to calculate and treat analytically, and (iv) it seemed equally justified to us for small sizes because very little is known about the true carrying capacity of a tumor microenvironment in its early stages. Lastly, the number of new drivers δn_d and passengers δn_p acquired during cell division are Poisson-distributed random variables with mean μ_d and μ_p , respectively, i.e. $P(\delta n_d = k|\mu_d) = \frac{(\mu_d)^k e^{-\mu_d}}{k!}$ and $P(\delta n_p = k|\mu_p) = \frac{(\mu_p)^k e^{-\mu_p}}{k!}$. In this model, mutations arise at a rate proportional to generation time, rather than absolute time. Our choice, therefore, reflects our belief that most mutational processes occur in a cell division-dependent manner (e.g. mutations in cell-cycle checkpoints like p53), rather than independent of cell division. Certainly, some mutational processes occur independent of cell division. However, because the generation time in precancerous cells accelerates only mildly over the course of progression, introducing mutations at an absolute rate would not significantly alter dynamics.

Many of the particular design choices and properties of our model were altered and then investigated in a previous study [2]. Specifically, we considered (1) the effects of mutations with additive effects [i.e. $B(n_d, n_p) = 1 + n_d s_d - n_p s_p$], (2) the effects of mutations that alter the death rate [i.e. $D = D(N, n_d, n_p)$], (3) the effects of driver and passenger mutations selected from various distributions (exponential, normal, and gamma) of fitness effect sizes, and (4) variations on the relation between population size and death rate. For the parameters that we believe are most relevant to cancer (**Table S1**), these permutations did not qualitatively alter our simulations. However, in the analytical analysis presented below we discuss the boundaries where assumptions of our model break-down; this was, in part, why we analyzed the model in such detail.

Many of the considerations discussed above are germane to all models of tumor progression, not simply the *in silico* model presented here. Consider that recent data on growth rates of human tumors differs from data obtain from mouse models: human tumors grow according to an exponential curve [4], while mouse tumors grow according to a Gompertz curve [5]. Careful mathematical consideration of the differences between a model of progression where growth is exponential, and one where growth is Gompertzian, should allow us to under-

stand when it is necessary to refine the design of simulations and experiments or change our conclusions.

The mutation rate not only varies considerably between tumors [6], but also varies as a tumor evolves. We believe that genomic instability happens early during progression, as this has been shown experimentally in some tumors [7] and suggested to be the first event during progression by virtue of theoretical considerations [8]. However this presumably differs from tumor to tumor. Thus, by developing a theoretical understanding of the process, we hope to gain some intuitive understanding of how variation in mutation rate over time might alter dynamics. For example, we show later that for tumors far below the critical mutation rate μ^* , variation in the mutation rate increases or decreases the rate of accumulation of drivers and passengers equally. However, near or above μ^* , variation in the mutation rate has a profound impact on driver's probability of fixation and effect size.

Before describing the entire dynamics of our model, it is useful to consider the difference between our simulations initiated at their stationary size (N^0 cells) and simulations initiated at 1 cell. In the absence of mutations, an initial population of one cell will grow logistically until it reaches the stationary size. Hence, it takes approximately $\text{Log}_2[N^0] \sim \text{Log}_2[10^3] \sim 10$ generations for the initial cell to approach stationary size. This is far shorter than the average time required for cancer progression ($\sim 10,000$ generations) and the time required for a new driver to accumulate ($\sim 1/(\mu_d N^0 s_d) \sim 1,000$ generations). Thus, our choice of initiating a tumor at one cell versus N^0 does not significantly alter the conclusions of our model.

This comparison of timescales also suggests that cancers are almost always near their stationary size:

$$\overline{B(n_d, n_p)} \approx D(N)$$

We previously tested this conclusion in simulations and found that it is an excellent approximation of tumor size [2]. If we assume $\overline{B(n_d, n_p)} \approx B(\overline{n_d}, \overline{n_p})$, then a relationship between the number of drivers and passengers in a tumor and its size is obtainable:

$$\begin{aligned} B(\overline{n_d}, \overline{n_p}) &\approx D(N) \\ \frac{(1 + s_d)^{\overline{n_d}}}{(1 + s_p)^{\overline{n_p}}} &\approx \text{Log}\left[1 + \frac{N}{(e - 1)N^0}\right] \\ \overline{n_d}\text{Log}(1 + s_d) - \overline{n_p}\text{Log}(1 + s_p) &\approx \text{Log}\left[\text{Log}\left[\frac{N}{N^0}\right]\right] \end{aligned}$$

Hence, if $s_d, s_p \ll 1$, then:

$$\begin{aligned} \overline{n_d}s_d - \overline{n_p}s_p &\approx \text{Log}\left[\text{Log}\left[\frac{N}{N^0}\right]\right] \\ &\approx \text{Log}[D(N)] \end{aligned} \tag{S3}$$

This final equation suggests that there exist a linear relationship between drivers and passengers among tumors with similar s_d and s_p , which we assume is the

case for tumors of the same tissue of origin. The relationship should be relatively robust to tumor size, but sensitive to the fitness effects of drivers and passengers. Moreover, changes in the functional form of $D(N)$ will alter the y-intercept of this linear relationship, but not the slope of the relationship. Hence, we can draw conclusions about the relative strength of drivers versus passengers (s_d/s_p) without knowing the exact constraints on population size. We tested and verified this prediction of a linear relationship between drivers and passengers in the main text.

1.2 Parameter range investigated

Our computational model has 5 independent parameters: a mutation rate μ , a mutation’s relative likelihood of being a driver versus a passenger T_d/T_p , the fitness benefit of a driver s_d , the fitness disadvantage of a passenger s_p , and an initial stationary size N_0 . These parameters vary considerably between tumor types (and the mutation rate even varies within tumor types [6]), so we explored a wide range of values centered around literature best-estimates (**Table S1**). More importantly, our analytical analysis reveals that we can describe our system with two dimensionless parameters, which we then estimated from age-incidence and genomics data (**Fig. 2**).

The most critical constraint of our parameters exploration is that $T_d \ll T_p$. Without this property a barrier to adaptation is not observed and infinite mutation rates become optimal. This constraint on target sizes for simulations is justified for a number of reasons. A priori, it should be expected that deleterious mutations outnumber advantageous mutations in natural populations simply because natural selection optimizes genomes to their environment—implying that most changes will be neutral or damaging. Indeed, most protein coding mutations and alterations were deleterious or neutral when investigated empirically in fly [9], yeast [10], and bacterial genomes [11]. We consider only moderately deleterious loci here ($s_p \approx 10^{-4} - 10^{-1}$), which nevertheless account for most nonsynonymous mutations [12,13]. Deleterious mutations outside of this range either do not fixate or negligibly alter progression [2].

There is also considerable evidence that $T_p \gg T_d$ in cancer (like natural populations). As much as 10% of the human genome is well-conserved and likely deleterious when mutated [14,15]. Conversely, there are only approximately 100-200 potential driver genes [16,17]. If driver loci include only a few specific sites per gene ~ 10 , then collectively drivers will constitute less than one one-millionth of the genome. Also, accumulated passengers greatly outnumber accumulated drivers (**Table S2**). This implies that the target size of passengers greatly outnumbers the target size of drivers, as selection can only increase the frequency of advantageous mutations relative to deleterious mutations.

For most of our parameter range $s_d > s_p$, however we do explore exceptions to this rule and our analytical model continues to explain dynamics well in these cases. The selection coefficients of drivers s_d and passengers s_p were estimated from genomics data in the main text and found to be comfortably within the range we explored. Nevertheless, there was good evidence for the range of fitness

benefits for drivers before we began our study. A previous study found that an s_d of 0.1 is necessary to obtain waiting times to cancer consistent with age-incidence rates [18]. In the main text we discuss evidence from mice models that support this approximate value of s_d .

Two previous investigations of cancer progression have considered deleterious passengers and found that they have a minimal impact on progression [19, 20]. The first paper assumes that passengers are effectively lethal to cancer cells (i.e. $s_p \rightarrow \infty$). They conclude that deleterious passengers are unimportant because they are quickly weeded out of the population. This is consistent with our results (see [2], or **Fig. S4**), however we believe that our best-estimate of $s_p \approx 10^{-3}$ is more reasonable. Our justification for this is discussed above, but also supported by two analyses that we performed in a previous study [2]. In this study, we observe very little negative selection in passenger genes (see Figure S6), which indicates that mutations in these genes cannot be lethal as they would never accumulate. We also found that mutations continue to accumulate even in regions classified as ‘deleterious’ within housekeeping genes (see Figure 4 of [2]), so even these very harmful mutations continue to fixate. To us, this is clear evidence that s_p cannot be large enough to prevent the fixation and accumulation of most deleterious passengers.

In a second study that concludes that deleterious passengers have a minimal impact on cancer progression [20], the authors assume that there exist only ~ 100 housekeeping genes in cancer that are deleterious when mutated. Again, we observe similar behavior in our simulations: when $T_p \approx T_d$, passengers do not appreciably alter progression. However, our best-estimate of the number of relevant deleterious genes is $50\times$ larger than their estimate (**Table S1**), while the paper discussed in the preceding paragraph argues that T_p is $100\times$ greater than this estimate [19]. There are two reasons for this discrepancy in parameter choice. First, [20] considers only deleterious housekeeping genes, while we believe many other genes and non-coding sequences (e.g. regulatory DNA sequences or microRNAs) could potentially be deleterious to cancer cells. Second, we find that the number of reported housekeeping genes is much larger than 100 (3,804 genes are classified as ‘housekeeping’ in [21], which also reviews other similar estimates). We believe most genes could be deleterious when mutated because (i) it has been proposed that passengers might invoke an immune reaction to tumor cells [22], and because (ii) passengers can cause cytotoxicity via protein imbalance and aggregation [23]. This later mode of damage should be applicable to nearly all expressed genes within the tumor, which constitute more than half of the 26,588 identified genes in the human genome [21].

Lastly, [20] remarks that passenger’s effect on dynamics weakens as s_p increases beyond 0.01. Again, this is consistent with our analyses, given the author’s small choice of T_p , but not consistent with results using our best-estimate of T_p . In our theoretical work below, we show that the optimum s_p for slowing progression s_p^* is approximately proportional to the mutation rate of passengers ($s_p^* \sim \mu_p$). Hence, for the authors chosen target size of passengers, their finding is correct; however, given our estimate of the number of deleterious passenger genes, increases of s_p beyond 0.01 continue to increase the drag of passengers

on cancer progression.

Because the relevance of deleterious passengers depends upon evolutionary parameters, their effects may be neglected in certain circumstances. In **Figure S4**, we identify the evolutionary regimes where passengers dominate, where they compete with drivers, and where they can be probably be neglected. We believe and present evidence that passengers are relevant for progression in most carcinomas, but we also observe that lymphomas fixate very few passengers [2]. Hence, it is important to consider the quantity of passengers accumulating in a tumor type before concluding that deleterious passenger can or cannot be neglected in the tumor type. All these concerns underscore the importance of further investigating the evolutionary parameters of cancer progression for various tumor types.

1.3 A two-hit model of cancer progression

It has been proposed that driver mutations may only be beneficial in a certain genetic context [24]. Oncogenes like c-Myc and k-Ras have been shown to induce senescence in some cancer cell lines unless they are accompanied by mutations in p53 or other associated proteins [25]. Likewise, many tumor suppressors mutations are recessive and require a second 'Loss of Heterozygosity' (LOH) event to impart their phenotypic effects. These types of mutations have been described as operating via a 'two-hit' process: the first driver event confers no change to cell fitness, while the second genetic event confers the benefit of both mutations.

Two-hit models have been studied previously in cancer evolution ([3, 8, 26] to name a few), but never in the presence of frequent moderately-deleterious passenger mutations. For this reason, we considered our originally defined model above, modified such that the first driver mutation confers no benefit to the cell, while the second driver mutation confers the benefit of both mutations, and all remaining mutations confer a benefit of s_d . Hence,

$$w(n_d, n_p) = \begin{cases} (1 + s_p)^{n_p}, & \text{if } n_d < 2 \\ (1 + s_d)^{n_d} (1 + s_p)^{n_p}, & \text{otherwise} \end{cases}$$

We kept all other properties of the model the same and investigated the result of this permutation in **Figure S7**. We observe three changes:

1. While the critical barrier to progression remains, its location N^* increases;
2. Initial trajectories slowly decay in a long period of stasis that allows additional passengers to accrue and delays progression;
3. The transition from the non-adaptive to adaptive regime is slower, as the period of stasis is highly variable.

We can understand these observations by first considering the mean time of stasis (time until the second driver mutation fixates) for the population. The

probability that a second driver fixates in the population at time t is simply the probability that a cell already harbors the first driver mutation ($\mu_d t$), times the total number of cells in the population ($N_0 e^{-v_p t}$), times the probability that the second driver arises and sweeps through the population ($\mu_d \frac{2s_d}{1+2s_d}$). Hence, the probability that stasis will end at generation t is $P_{\text{exit}}(t) = N_0 \mu_d^2 \frac{2s_d}{1+2s_d} t e^{-v_p t}$. Thus, the probability that stasis last a certain number of generations $P_{\text{stasis}}(t)$ is the probability of not exiting in the prior generations and also exiting at generation t :

$$\begin{aligned} dP_{\text{stasis}}(t) &= -P_{\text{exit}}(t) dt P_{\text{stasis}}(t) \\ P_{\text{stasis}}(t) &= e^{-\int_0^t P_{\text{exit}}(t') dt'} \end{aligned} \quad [\text{S4}]$$

From simulations (**Fig. S7**), it is clear that most simulations that exit this early stasis period go onto progress to cancer. Hence to a first approximation, the new $N^{*,\text{two-hit}}$ is the value of N_0 , where $P_{\text{stasis}} = 50\%$, thus:

$$\begin{aligned} \text{Log}(2) &= \int_0^\infty N^{*,\text{two-hit}} \mu_d^2 \frac{2s_d}{1+2s_d} t e^{-v_p t} dt \\ N^{*,\text{two-hit}} &= \frac{\text{Log}(2)(1+s_d)v_p^2}{2\mu_d^2 s_d} \end{aligned}$$

This predicts $N^{*,\text{two-hit}}$ to be 883 for the trajectories plotted in **Figure S7**, which is within a factor of two from the 50% success point. It is somewhat below the observed tradition point presumably because trajectories that exit this stasis period far below N^* never progress, and because segregating passengers may interfere with the first beneficial driver. Collectively, these results suggest that our mathematical framework of our model is generally applicable to cancers where the first driver isn't beneficial, but that there are also some dramatic differences that warrant further investigation.

While two-hit models of progression are most likely applicable in many tumors, we do not believe these models are more universal than our original formulation. Consider that in experiments directly measuring the change in cell fitness upon activating mutations in k-Ras, an immediate increase in proliferation was observed [27]. Also, many tumor suppressors are haploinsufficient (e.g. Dicer [28], p27 [29], CDC4 [30], p18 [31]). These results suggest that driver mutations are often advantageous, after only one mutation, if perhaps at a muted level.

2 An analytical model of dynamics

In the main text, we demonstrate that dynamics are described by two counteracting forces: an upward velocity v_d resulting from accumulating beneficial drivers, and a downward velocity v_p resulting from accumulating deleterious passengers. The upward velocity v_d was further subdivided into a product of the rate at which new drivers fixate in the population f times their effect on population size once fixated ΔN (**Fig. 1B**)². The velocities v_d and v_p are bal-

²While we assume that drivers arise at random time intervals, this assumption is not always true. Because unfixed passengers can interfere with the fixation of drivers, a driver is

anced at a critical population size N^* , at which the population is approximately equally likely to go extinct or progress to cancer.

While we were able to describe the average behavior of our population in the main text, our model (like cancer) is inherently stochastic. The complete dynamics of our model are best described by a differential equation with stochastic jumps:

$$\begin{aligned} dN &= v_p dt + \Delta N dn_d \\ n_d &\xrightarrow{f} n_d + 1 \end{aligned} \tag{S5}$$

In this equation, the change in population size dN is the product of a deterministic component $-v_p$, along with a stochastic component describing the random arrival of new drivers ($\Delta N dn_d$). Below, we use this equation to estimate the probability of cancer for any population size $P_{\text{cancer}}(x)$ and the mean waiting time to cancer $\overline{t_{\text{cancer}}}(x)$, where $x = N/N^0$ is a dimensionless population size discussed in the main text and below. Lastly, we noticed that simulations differed from the formalism we presented in the main text when we varied the mutation rate μ and explored a broader range of passenger deleteriousness s_p (**Fig. 2, S4**). These discrepancies could be resolved by considering two phenomena neglected by our first derivation: selection against passengers, and passenger’s effect on both the fixation probability and clone fitness of drivers. Fortunately, accounting for these phenomena did not alter Eq. 5, nor the overall framework of our analytical model. Instead, they only affect the rates v_p , f , and jump size ΔN in our model. Thus, with the refined formalism, we described dynamics across a very broad range of parameters (**Fig. 2, S4**). More importantly, we observe drastic reductions in the probability of adaptation at high mutation rates and when passengers are moderately deleterious. These findings suggest novel strategies to cancer therapy.

From Eq. 5, it is evident that population size is the state variable of our system and, as such, is all that is needed to describe future dynamics (this is also observed in **Fig. S5**). By converting population size into a dimensionless parameter $x = N/N^*$ (and $x^0 = N^0/N^*$), the probability of cancer collapse onto a simple curve $P_{\text{cancer}}(x)$ (**Fig. 1**)—further underscoring the importance of the critical population size. Hence, we will use this dimensionless quantity x heavily throughout the remainder of our analysis.

2.1 Estimating the probability of cancer

Using Eq. 5 we can describe how the probability of extinction changes in an infinitesimal time due to either passenger accumulation or a rare driver jump:

$$P_{\text{cancer}}(x) = f(x)dtP_{\text{cancer}}[x + \Delta N(x)] + [1 - f(x)dt]P_{\text{cancer}}[x - v_p(x)dt]$$

In this equation, we see that is the probability of cancer at x is the probability of a jump times the probability of cancer after the jump ($f(x)dtP_{\text{cancer}}[x +$

more likely to fixate immediately following a previous driver fixation event [32]. Ignoring this caveat does not significantly alter dynamics in the parameter space explored here.

$\Delta N(x)$) plus the probability of decline times the probability of cancer after the decline ($[1 - f(x)dt]P_{\text{cancer}}[x - v_p(x)dt]$). Note that f , ΔN , and v_p are all functions of x in the equation above. Defining these functions in such a general form makes solving the stochastic differential equation impossible. So we note that each function is approximately linear in x . Thus, we can replace each function with a constant times x : $f(x) \rightarrow fx$, $\Delta N(x) \rightarrow \Delta Nx$, and $v_p(x) \rightarrow v_px$. The probability of cancer after a decline can be expanded via a Taylor series: $P_{\text{cancer}}[x - v_pxdt] \approx P_{\text{cancer}}(x) - v_pxdtP'_{\text{cancer}}(x)$. Along with the linear approximations for f and θ , this reduces the above equation to:

$$\frac{v_p}{f} P'_{\text{cancer}}(x) = P_{\text{cancer}}(\theta x) - P_{\text{cancer}}(x) \quad [\text{S6}]$$

Here, $\theta = 1 + \Delta N \approx 1 + s_d$ denotes the logarithmic change in population size after a driver jump. Next we notice from simulations that $P_{\text{cancer}}(x)$ changes most significantly when $x \approx 1$. Hence, we can logarithmically-transform x and solve this new variable $y = \log(x)$ via a Maclaurin Series:

$$\begin{aligned} \frac{v_p}{f} \frac{dP_{\text{cancer}}(y)}{dy} &= e^y(P(y + \text{Log}(\theta)) - P(y)) \\ &\approx e^y(\text{Log}(\theta) \frac{dP_{\text{cancer}}(y)}{dy} + \frac{1}{2} \text{Log}^2(\theta) \frac{d^2P_{\text{cancer}}(y)}{dy^2} + \dots) \end{aligned}$$

Now, by reverting from y back to x , we obtain:

$$\begin{aligned} \frac{v_p}{f} P'_{\text{cancer}}(x) &= \\ &\text{Log}(\theta)xP'_{\text{cancer}}(x) + \frac{1}{2}\text{Log}^2(\theta)x^2P''_{\text{cancer}}(x) + \frac{1}{2}\text{Log}^2(\theta)xP'_{\text{cancer}}(x) \quad [\text{S7}] \end{aligned}$$

By eliminating the last term in this solution, a reasonable approximation because $\text{Log}^2(\theta) \ll 1$, the differential equation is now solvable. Its boundary conditions (essentially the definitions of *cancer* and *extinction*) are:

$$\begin{aligned} P_{\text{cancer}}(x = 0) &= 0 \\ P_{\text{cancer}}(x = \infty) &= 1 \end{aligned}$$

The probability of cancer after infinite time (demonstrable by substitution into Eq. 7):

$$P_{\text{cancer}} = 1 - \gamma\left(\frac{2}{\text{Log}(\theta)}, \frac{2}{\text{Log}(\theta)x}\right) \quad [\text{S8}]$$

Here, $\gamma(s, x) = 1/\Gamma(s) \int_0^x e^{-t}t^{s-1}dt$; $\Gamma(s) = \int_0^\infty x^{s-1}e^{-x}dx$ is the normalized incomplete gamma function. This solution is parameterized by two dimensionless quantities: θ and x , which represent the jump size in population of driver sweeps and our effective population size respectively.

2.2 Estimating the mean time to progression

We can also use Eq. 5 to solve for the waiting time to cancer. This can be accomplished in two ways: (1) we can simulate random driver jumps and deterministic passenger decline directly, and (2) we can approximate the mean waiting time to cancer using a Taylor expansion similar to the strategy we employed to solve for the probability of cancer. These two approaches agree with each other (thus, illustrating their accuracy), and offer key insights into the evolutionary parameters that affect age-incidence curves (**Fig. 2, S1**).

Eq. 5 can be simulated using a “hybrid” Gillespie algorithm: a meta-simulation of driver- and passenger-accumulation events that we, originally, observed arising from our atomistic simulations of birth, death, and mutational events. The advantage of this technique is that it allows us to quickly simulate billions of tumors, which would be computationally impossible via full-detail simulations. Because we are confident that we are accurately estimating the rate of driver and passenger accumulation events (**Fig. S4**), this simplification should retain accuracy. To simulate Eq. 5 directly, we must consider that the instantaneous probability of a driver jump $f[x(t)]$ is a function of a constantly declining population size due to passenger accumulation: $x(t) = x_{n_d}(1 + s_p)^{v_p/s_p t} \approx x_{n_d}e^{-v_p t}$. Here, x_{n_d} is the population size after the last driver jump. Thus, the waiting time between drivers $\Delta t = t_{n_d+1} - t_{n_d}$ is:

$$\begin{aligned} \int_0^{\Delta t} fN(t')dt' &= \zeta \\ f \int_0^{\Delta t} N_{n_d}e^{-v_p t'} dt' &= \zeta \\ \Delta t &= -\frac{1}{v_p} \text{Log}\left(1 - \frac{v_p \zeta}{fN_{n_d}}\right) \end{aligned} \quad [\text{S9}]$$

ζ is an exponentially-distributed random number with mean 1. Using our precise calculations of f , v_p and ΔN below, we can now simulate Eq. 5 directly.

We can also solve Eq. 5 for t_{cancer} , using the exact same approximations as we did to estimate $P_{\text{cancer}}(x)$. To do this, we begin with a Master Equation for the probability of acquiring a cancer after waiting time t when currently at size x :

$$P_{\text{cancer}}(x, t) = f(x)\delta t P_{\text{cancer}}(\theta x, t + \delta t) + [1 - f(x)\delta t]P_{\text{cancer}}[x - v_p(x)\delta t, t + \delta t]$$

The mean waiting time to cancer is then:

$$\overline{t_{\text{cancer}}(x)} = \int_0^{\infty} t P_{\text{cancer}}(x, t) dt$$

Before substituting the Master Equation into this definition, we must first utilize a first-order Taylor series expansion about θ and δt :

$$P_{\text{cancer}}(\theta x, t - \delta t) \approx P_{\text{cancer}}(x, t) + \frac{\partial P_{\text{cancer}}(x, t)}{\partial x}(\theta - 1) + \frac{\partial P_{\text{cancer}}(x, t)}{\partial t}\delta t$$

This leads to the solution:

$$\begin{aligned}
\overline{t_{\text{cancer}}}(x) &= \int_0^\infty P_{\text{cancer}}(x, t) t dt \\
&\quad + \delta t \int_0^\infty \frac{\partial P_{\text{cancer}}(x, t)}{\partial t} t dt \\
&\quad + [f \delta t (\theta - 1) + (1 - f)(v_p \delta t)] \int_0^\infty \frac{\partial P_{\text{cancer}}(x, t)}{\partial x} t dt
\end{aligned}$$

The first integral in this solution is simply the definition of our mean waiting time ($\overline{t_{\text{cancer}}}(x)$). The second integral can be integrated by parts by noting that $\lim_{t \rightarrow \infty} t P_{\text{cancer}}(x, t) = 0$ (otherwise, $\overline{t_{\text{cancer}}}(x)$ would be undefined). Lastly, the third integral reduces to $t'_{\text{cancer}}(x)$. Thus, we eventually find:

$$f(x)[\rho_c(\theta x) - \rho_c(x)] - v_p(x)\rho'_c(x) + P_{\text{cancer}}(x) = 0$$

Here, $\rho_c(x) = P_{\text{cancer}}(x)\overline{t_{\text{cancer}}}(x)$. This equation has a nearly identical form to Eq. 6. So we used a similar Second-Order Maclaurin series expansion of $\text{Log}(x)$ to approximate its solution:

$$\begin{aligned}
\overline{t_{\text{cancer}}}(x) &= \frac{2}{f \text{Log}^2(\theta)} \left[\int_x^\infty \frac{dy}{y^3} \frac{P_{\text{cancer}}(y)[1 - P_{\text{cancer}}(y)]}{P'_{\text{cancer}}(y)} \right. \\
&\quad \left. + \frac{1 - P_{\text{cancer}}(x)}{P_{\text{cancer}}(x)} \int_0^x \frac{dy}{y^3} \frac{P_{\text{cancer}}^2(y)}{P'_{\text{cancer}}(y)} \right] \quad [\text{S10}]
\end{aligned}$$

These integrals can be numerically computed using Simpson's Method and yield a solution that is in good agreement with the hybrid simulations described in the preceding paragraph (**Fig. S1**).

Our solution for the waiting time to cancer is most illustrative when $x \ll 1$ —the regime that we expect to contain most tumors. In this regime, the mean time of cancer progression increases as $-\text{Log}(x)/v_p$, which implies two interesting properties of $\overline{t_{\text{cancer}}}$. First, x has a very weak, sub-linear, effect on the waiting time and does not significantly alter the shape of incidence curves (**Fig. S1**). Second, the waiting time to cancer is dictated by v_p (the accumulation rate of passengers), thus offering yet another reason to continue investigating the rate of deleterious passenger accumulation.

The mean time to cancer $\overline{t_{\text{cancer}}}(x)$ is a quantity that is conditioned on a population actually progressing to cancer. Hence, it depends heavily on the probability of adaptation $P_{\text{cancer}}(x)$. Because $P_{\text{cancer}}(x)$ has an inflection point at $x = 1$, $\overline{t_{\text{cancer}}}(x)$ behaves very differently when $x > 1$, than when $x < 1$. When $x > 1$, $\overline{t_{\text{cancer}}} \approx \int \frac{1}{\langle dN/dt \rangle} dN$ (i.e. the waiting time to cancer is what would be expected from our mean-velocity formulation), as nearly all cancers succeed. However, when $x < 1$, $\overline{t_{\text{cancer}}}(x)$ defies mean behavior—the average cancer goes extinct. Only the rare, exceptional populations that progress to cancer are weighted in the mean of $\overline{t_{\text{cancer}}}(x)$; these exceptional populations grow much faster than the average population. Hence, the increase in waiting time to cancer grows sub-linearly with x , when $x < 1$.

2.3 Accumulation of deleterious passengers

Passenger mutations accumulate and drag populations down with a rate v_p . This quantity is a product of passenger’s arrival rate $\mu_p N$, their fixation probability π_p , and their effect on population size once fixated $N s_p$ (i.e. $v_p \approx \mu_p s_p N$). In the main text, we assume that the fixation probability is approximately neutral ($\pi_p \approx 1/N$); however, when selection is stronger than genetic drift, the fixation probability becomes less than the neutral rate. A number of studies have focused on the fixation probability of deleterious mutations in a population, termed Muller’s Ratchet [33–36]. In general, estimates of Muller’s Ratchet (and consequentially π_p) begin by considering the distribution of deleterious alleles in a population of infinite size in mutation-selection balance—where allele frequencies are not changing. At equilibrium, such a population exhibits a Poisson distribution in the number of segregating passengers δp within cells $N_{\delta p}$, defined by a characteristic parameter $\lambda_p = \mu_p/s_p$ (**Fig. 3C**):

$$N_{\delta p} = N \frac{e^{-\lambda_p} \lambda_p^{\delta p}}{\delta p!} \quad [\text{S11}]$$

If we then consider a population of finite size, we find that the allele frequencies fluctuate due to genetic drift. If fluctuations in the fittest class ($N_{\delta p=0} = N e^{-\lambda_p}$) are large enough to cause this fittest class to go extinct, then it is irrevocably lost from the population. This irrevocable loss is considered a ‘click’ of Muller’s Ratchet. The new fittest class—individuals harboring one segregating passenger prior to the ‘click’—then relaxes to a new equilibrium that fluctuates, and the process repeats. Estimating the time required for a new fittest class to relax to equilibrium size immediately following a ‘click’ is non-trivial and dependent upon the parameters of the system: N , s_p , and μ_p , which can vary by orders of magnitude depending upon the evolutionary system in question; hence there are many estimates of the exact rate of Muller’s Ratchet.

We present and utilize 3 estimates of the rate of Muller’s Ratchet:

1. A solution that works well for most values of s_p , μ_p , and N considered here (**Fig. 2, magenta lines**) and simply ignores the time to equilibration after ‘clicks’;
2. A traveling-wave solution accurate for large values of λ_p [35] that allows the distribution of segregating passengers to be far from equilibrium, but presumes that the size of neighboring mutational classes are uncorrelated; and
3. A solution accurate for small values of λ_p [36] that considers correlations between neighboring fitness classes, but requires that the population be in quasi-equilibrium (i.e. near mutation-selection balance).

Estimates (2) and (3) accurately describe Muller’s Ratchet across complementary regions of our phase space. By combining these later two estimates with

estimates of the number of hitchhiking passengers and their effects on the probability of driver fixation events, we developed a precise description of our model's dynamics (**Fig. 3, S4; black lines**).

If we simply ignore the time required for a population to equilibrate into mutation-selection balance, then we can estimate the rate of Muller's Ratchet with a closed form solution that is applicable to all values of s_p , μ_p , and N investigated here. We assume that the probability of a 'click' is approximately the probability of a new passenger fixating *within* the fittest class: $N_{\delta p=0} = Ne^{-\lambda_p}$. In other words, to a first-approximation, deleterious passengers simply reduce the effective population size of our system, such that $N_e \sim Ne^{-\lambda_p}$. The probability of a lone deleterious allele fixating within this fittest class is describe by a Moran Process [37]. Hence,

$$\pi_p^{(1)} = \frac{s_p}{(1 + s_p)^{N_e} - 1} \quad [\text{S12}]$$

This refined fixation probability $\pi_p^{(1)}$ is then used to correct the downward velocity due to passengers, using the same formula for v_p derived in the main text:

$$v_p^{(i)} = \mu_p s_p N \pi_p^{(i)} \quad [\text{S13}]$$

This equation links v_p to the passenger fixation probabilities calculated above, and the other two fixation probabilities calculated below.

The solution for Muller's Ratchet as a traveling wave, which we apply when $\lambda_p < 1$, was obtained from [35]:

$$\frac{\text{Log}\left(\frac{Ns_p}{\sqrt{\lambda_p}}\right)}{\lambda_p} \approx 1 - \frac{\pi_p^{(2)}}{2} \left[\text{Log}^2\left(\frac{e}{\pi_p^{(2)}}\right) + 1 \right] - \frac{1}{\lambda_p} \text{Log} \left[\frac{(\pi_p^{(2)})^{3/2}}{\sqrt{1 - p i_p^{(2)}}} \frac{\text{Log}\left(\frac{e}{\lambda_p}\right)}{1 - \pi_p^{(2)} \text{Log}\left(\frac{e}{\lambda_p}\right) + \frac{5}{6\lambda_p}} \right] \quad [\text{S14}]$$

Because this equation is transcendental, we solved for $\pi_p^{(2)}$ using Brent's Method.

When $\lambda_p \geq 1$, a quasi-stationary analysis of the mutation classes becomes appropriate. This analysis was first done in [36], resulting in a solution of the form:

$$T_{click} = \frac{e - 1}{s_p} e^{\frac{s_p N p_0}{2(e-1)}} \quad [\text{S15}]$$

The fixation probability is then simply the inverse of the 'click' time: $\pi_p^{(3)} = 1/T_{click}$.

Lastly, there is a discontinuity between the above two solutions at their intersection: $\lambda_p = 1$. We resolved this by interpolating between the two solutions, as follows:

$$\pi_p^{(\text{combined})} = \lambda_p \pi_p^{(2)} + (1 - \lambda_p) \pi^{(3)}$$

2.4 Effects of deleterious passengers on fixation probability and clone fitness of drivers

The occurrence and fixation of driver mutations are rare events, separated by nearly random time intervals, with a frequency of occurrence $f = \mu_d N \pi_d$. Here, π_d is the fixation probability of a new mutant driver once it arises in the population. In the first-order model presented in the main text, we estimate that $\pi_d = s_d / (1 + s_d) \approx s_d$. However, this result assumes that there are no other non-neutral alleles in the population. In reality, there are many segregating passengers in the population, and potentially other segregating drivers.

The presence of other drivers in the population, which interfere with the fixation of our clone of interest, is a phenomena commonly described as *Clonal Interference* [38]. Clonal Interference becomes significant in the population once the time required for a driver to fixate [$\sim \text{Log}(N)/s_d$ generations] approaches the fixation rate ($f \approx \mu_d N s_d$). Nascent precancerous population are in a space of evolutionary parameters where Clonal Interference is particularly negligible: population size is small ($N \sim 10^3$), and drivers are rare ($\mu_d \sim 10^{-5}$), but strong ($s_d \sim 10^{-1}$). Thus, we do not consider its effects here. However, for a larger tumor population, clonal interference may become very significant. This is especially true in a poorly-mixed population, where beneficial alleles take longer to sweep through the population [39].

Segregating passenger mutations can also interfere with a driver sweep by ‘hitchhiking’ on the expanding clone [32, 40]. Most of the analysis we present here has already been presented in these two previous works, but because it is integral to our derivation of the critical mutation rate and because we extend their analysis, we have decided to repeat their work.

For mathematical analysis, we disentangle two types of hitchhikers: (1) those that reside in the Initial clone before the new driver arises (denoted δp_I), and (2) those that arise and fixate in the new driver clone as it Sweeps through the population (denoted δp_S). It is necessary to distinguish hitchhikers this way because only the initial hitchhikers (δp_I) significantly alter the fixation probability f , while both types alter the effect size ΔN . The hitchhikers that accumulate during the sweep will generally arise after the clone is of appreciable size; however, once the driver clone is of appreciable size, it is exceedingly likely that it will fixate so long as it remains the fittest clone in the population.

Here, we consider only the average number of hitchhikers in a driver sweep ($\overline{\delta p_I}$ and $\overline{\delta p_S}$), rather than their entire distribution of quantities; estimates of the average number of hitchhikers appear to explain dynamics reasonably well (first shown in [32] and also evident from our analysis’ good agreement with simulations **Fig. S4**). Thus the probability that a new clone fixates in the absence of Clonal Interference is (**Fig. 3C**):

$$\pi_d(\overline{\delta p_I}) = \frac{s'_d(\overline{\delta p_I})}{1 + s'_d(\overline{\delta p_I})} : s'_d(\overline{\delta p_I}) = s_d - \overline{\delta p_I} s_p \quad [\text{S16}]$$

The jump size ΔN becomes:

$$\Delta N' = N[s_d - (\overline{\delta p_I} + \overline{\delta p_S})s_p] \quad [\text{S17}]$$

We can conclude our analysis of hitchhikers once we obtain $\overline{\delta p_I}$ and $\overline{\delta p_S}$. These quantities were first derived in [32]. We use their results (summarized below), along with a minor necessary adjustment for populations when λ_p is large, to complete our analytical model of cancer progression.

For a new driver clone to take over the population and fixate, it has been shown that its fitness must be greater than the fittest class in the population [32]. This imposes a maximum on the number of initial hitchhikers δp_I^{\max} that a successful driver clone can have:

$$\begin{aligned} s_d &> \delta p_I s_p \\ \delta p_I^{\max} &= \lfloor s_d/s_p \rfloor \end{aligned}$$

A clone that does not satisfy this constraint may proliferate for a while in the population, but it will nevertheless be eventually out-competed by fitter clones. When the mean number of hitchhiking passengers (λ_p) approaches this maximum, hitchhikers dramatically reduce both f and ΔN , thus increasing N^* to untenable sizes. This occurs when:

$$\begin{aligned} \lambda_p &= \delta p_I^{\max} \\ \mu_p/s_p &= \lfloor s_d/s_p \rfloor \\ \mu_p &\approx s_d \end{aligned} \quad [\text{S18}]$$

Hence, our analysis suggests a limit on the maximum mutation rate that an adapting population can tolerate: $\mu_p^* \approx s_d$. In simulations, we observe extinction slightly above this threshold (**Fig. 3A, S2**). This mechanism of collapse, where populations go extinct by failing to acquire new advantageous mutations or adaptations, differs from the traditional model of mutational meltdown. In the traditional model, advantageous mutations are generally ignored and meltdown occurs only because deleterious mutations accumulate too quickly. In our model, however, traditional mutational meltdown is difficult because populations also acquire advantageous mutations faster as the mutation rate increases. Moreover, traditional meltdown occurs only when the population size is small, making it impossible to occur in a large population like cancer. Our discovery of a new mechanism of meltdown that is independent of population size suggests that mutational meltdown may be induced via cancer therapeutics.

The number of initial segregating passengers in a clone when a driver arises (δp_I) can be obtained by considering, once again, the population at mutation selection balance, i.e. Eq. 11. The average number of initial hitchhiking passengers is simply the average of the likelihood of a driver arising in each mutational class, conditional on the driver successfully sweeping through the population:

$$\begin{aligned}
P(\delta p_I) &= \frac{1}{\mathcal{N}} N_{n_p = \delta p_I} \pi_d(\delta p_I) \\
\overline{\delta p_I} &= \frac{1}{\mathcal{N}} \sum_{\delta p_I=0}^{\delta p_I^{\max}} P(\delta p_I) \pi_d(\delta p_I) \\
&= \frac{1}{\mathcal{N}} \sum_{\delta p_I=0}^{\delta p_I^{\max}} \frac{e^{-\lambda_p} \lambda_p^{\delta p_I}}{\delta p_I!} \frac{s'_d(\delta p_I)}{1+s'_d(\delta p_I)}
\end{aligned} \tag{S19}$$

Here, $\mathcal{N} = \sum_{\delta p_I=0}^{\delta p_I^{\max}} \pi'_d(\delta p_I)$ is a normalization constant.

The above solution fails when λ_p is large. In this circumstance, the population is far from mutation-selection balance. Rectifying the solution in this case is difficult to do precisely, however a simple correction to Eq. 19 can crudely ameliorate the estimate. Because the assumption of mutation-selection balance fails only once the expected number of passengers in the fittest class becomes very small ($N_{n_p=0} = N e^{-\lambda_p} \sim 1$), we propose that the actual fittest surviving class in the population is the first class of passengers with an expected population size that is greater than the size of fluctuations in the population. Because the variance in a birth and death process is the sum of the rates ($2N$ in our model), the Fittest Surviving Class k_{FSC} is:

$$\begin{aligned}
k_{\text{FSC}} &= \min_{n_p} [N_{n_p} > \sqrt{2N}] \\
k_{\text{FSC}} &= \min_{n_p} [e^{-\lambda_p} \lambda_p^{n_p} / n_p! > \sqrt{\frac{2}{N}}]
\end{aligned}$$

The corrected distribution of δp_I then becomes:

$$\overline{\delta p_I} = \frac{1}{\mathcal{N}} \sum_{\delta p_I=0}^{\delta p_I^{\max}} P(k = \delta p_I + k_{\text{FSC}} | \lambda_p) \pi_d(\delta p_I)$$

This simple correct yields a final solution for P_{cancer} that agrees with simulations well (**Fig. S4**).

Lastly, the number of passengers that accumulate during the selective sweep (δp_S) can be calculated using a recursive relationship. This relationship begins with the probability of accumulating the maximum possible passengers during the sweep δp_I^{\max} [32]:

$$\begin{aligned}
P(\delta p_S = \delta p_I^{\max}) &= \frac{1}{\mathcal{N}_2} \delta p_I^{\max} \\
P(\delta p_S = k) &= \frac{1}{\mathcal{N}_2} \frac{k + s_p P(\delta p_S = k+1)}{1 + s_p} \\
\overline{\delta p_S} &= \frac{1}{\mathcal{N}_2} \sum_{\delta p_S=0}^{\delta p_I^{\max}} P(\delta p_S)
\end{aligned}$$

Where $\mathcal{N}_2 = \sum_{\delta p_S=0}^{\delta p_I^{\max}} P(\delta p_S)$ is a second normalization constant.

3 A traditional model of cancer progression with drivers and neutral passengers.

In the traditional model of cancer progression used to estimate age-incidence curves, it is assumed that a cancerous population transitions through k intermediate states before malignancy:

$$C_0 \xrightarrow{r_1} C_1 \xrightarrow{r_2} \dots \xrightarrow{r_k} C_k$$

Simply put, these intermediate states and transitions correspond to the many phenotypic changes that occur within a tumor as it progresses [41]. The instantaneous probabilities of each transition from one state to the next r_i can vary in the general case. Nevertheless, it has been shown that this predicts similar age-incidence rates to a model where transition rates are all the same [42]. Thus, for parsimony we only consider the case where all transition rates are the same constant r . Moreover, if the transition rates are drastically different from one another, then dynamics will largely be determined by the slowest rate alone. The faster rates are then no longer rate-limiting-steps and can be neglected.

From a genetic perspective, each transition corresponds to the acquisition of a new driver in the population. However from a mathematical perspective, this model is agnostic about the underlying molecular event that transitions a precancerous population from one state to the next. Thus, this model can be expanded to include any set of heritable rate-limiting steps required for carcinogenesis: SNMs, SCNAs, alterations in DNA and histone moieties, stable changes in cell signaling cascades, etc. Therefore, we believe it is reasonable to assume that each rate-limiting step is the acquisition of a new driver, as has been presumed for many years [43].

We now consider the properties of this model when neutral passengers (that do not alter progression) also accumulate. The precancerous population is now defined by the state C_{n_d, n_p} . We consider the case where drivers accumulate at a fixed rate r_d and passengers accumulate at different fixed rate r_p :

$$\begin{array}{ccccc}
 C_{0,0} & \xrightarrow{r_d} & C_{1,0} & \xrightarrow{r_d} & \dots \\
 \downarrow r_p & & \downarrow r_p & & \\
 C_{0,1} & \xrightarrow{r_d} & C_{1,1} & \xrightarrow{r_d} & \dots \\
 \downarrow r_p & & \downarrow r_p & & \\
 \vdots & & \vdots & \ddots & \\
 & & & & C_{n_d, n_p}
 \end{array}$$

As before, cancer arises once enough drivers accumulate ($C_{n_d=k, n_p}$).

To interpret age-incidence data, as well as genomics data, we are interested in both the waiting time until cancer (t_{cancer}) and the total number of mutations ($n_p + k$). This model can be simplify by noting that there is a freedom in the units for which we measure time. In our simulations, time was measured in generations and then converted to years. Here, we chose to measure time in units of the driver transition probability r_d and will then convert this to years afterwards. Hence, $r_d = 1$ without loss of generality. Consider the quantity $\tau_{\text{cancer}} = t_{\text{cancer}} r_d$, as a dimensionless measure of the waiting time to cancer. Its value will be roughly k on average. Because driver and passenger accumulation events are independent processes in this model, the joint probability of observing a cancer at time τ_{cancer} with n_p passenger mutations, $P(\tau_{\text{cancer}}, n_p | n_d = k, r_p)$, is:

$$P(\tau_{\text{cancer}}, n_p | n_d = k, r_p) = P(\tau_{\text{cancer}} | n_d = k) \cdot P(n_p | \tau_{\text{cancer}}, r_p) \quad [\text{S20}]$$

This joint probability distribution provides a framework for identifying our quantities of interest.

The waiting times to cancer in this neutral-passenger model, has been previously shown to be a sum of exponentially-distributed waiting times [42], i.e. an Erlang or Gamma distribution, of the form:

$$\begin{aligned} P(\tau_{\text{cancer}} | n_d = k) &= \text{Erlang}[\tau_{\text{cancer}} | n_d = k, r_d = 1] \\ &= \frac{r_d^k \tau_{\text{cancer}}^{k-1} e^{-r_d \tau_{\text{cancer}}}}{(k-1)!} \\ &= \tau_{\text{cancer}}^{k-1} e^{-\tau_{\text{cancer}}} / (k-1)! \\ &\propto t_{\text{cancer}}^{k-1}, \text{ when } \tau_{\text{cancer}}/k \ll 1 \end{aligned} \quad [\text{S21}]$$

Traditionally in this model, it is believed that very few precancerous population have enough time to progress, as lesion formation rates are much greater than cancer incidence rates. Hence, it is believed that age-incidence curves should be fit with only the beginning of this distribution: i.e. a power-law distribution (last line of Eq. 21). We find that although this hypothesis explains age-incidence rates well at mid-age, it fails to explain the plateau in age-incidence rates seen at older ages in most cancer subtypes (**Fig. 2A, S1**).

In this model, the total number of passengers accumulated is a Poisson distribution, if the time of progression t_{cancer} is known:

$$\begin{aligned} P(n_p | \tau_{\text{cancer}}, r_p) &= \text{Poisson}[n_p | \langle n_p \rangle = t_{\text{cancer}} r_p] \\ &= \frac{e^{-\langle n_p \rangle} \langle n_p \rangle^{n_p}}{n_p!} \end{aligned} \quad [\text{S22}]$$

Here, $\langle n_p \rangle = t_{\text{cancer}} r_p$ is the mean number of expected passengers. The distribution takes this form because each passenger accumulation event occurs with an exponentially-distributed waiting time, whose sum over a fixed time interval is described by a Poisson distribution. Because we do not know when a new lesion arrives, we must convolute this distribution with our expected distribution of t_{cancer} .

The available time for cancer progression depends upon the length of a human life: t_{human} . If $t_{\text{human}} < t_{\text{cancer}}$, then the precancerous population will be unobserved in age-incidence and genomics data because the person died of an alternate cause prior to malignancy. Although the actual distribution of human lifetimes is complicated, we can still make inferences about the validity of this model by considering its extremes. Consider two opposing extreme cases: (1) when $t_{\text{human}} \gg \overline{t_{\text{cancer}}}$, all lesions eventually progress and are sequenced (i.e. a human lifetime is much greater than the mean time to cancer); and (2) when $t_{\text{human}} \ll \overline{t_{\text{cancer}}}$, only a few exceptional lesions progress (i.e. the mean time to cancer is much shorter than a human lifetime). We find that this first extreme predicts a much broader and more positively skewed distribution in the number of passengers, than the second case. In either case, both distributions exhibit similar predicted distributions in the number of total mutations. We used the

second extreme where a human lifetime is much shorter than the mean time to progression, as this would be predicted by a power-law fit to age-incidence curves. This distribution is still not wide enough, nor positively skewed enough, to explain the observed distribution of passengers in cancer under realistic parameters (**Fig. 2B**, **Table S4**). In contrast, our model predicts a broader and positively skewed distribution that captures observed passenger histograms well (**Fig. 2B**).

In the case where $t_{\text{human}} \gg \overline{t_{\text{cancer}}}$, accumulation of passengers follows a binomial process. Each accumulation event has probability $p = r_d/(r_d + r_p)$ of being a driver and probability $(1 - p)$ of being a passenger. Because the population has infinite time to progress to cancer, the binomial process continues until $n_d = k$ drivers accumulate. A binomial process that continues until k successes (i.e. drivers), will have a total number of failures (i.e. passengers) that samples a negative binomial distribution:

$$P(n_p|p, k) = \binom{n_p + k - 1}{n_p} (1 - p)^{n_p} p^k \quad [\text{S23}]$$

A negative binomial distribution with $p \ll 1$ (i.e. passengers greatly outnumber drivers—as is the case in observed) reduces to a Poisson distribution.

In the case where $t_{\text{human}} \ll \overline{t_{\text{cancer}}}$, the waiting time to cancer follows a power law distribution (Eq. 21). This, convoluted with the distribution of passengers expected for a particular t_{cancer} (Eq. 22) yields the expected distribution of passengers for a cancer subtype:

$$\begin{aligned} P(n_p|k) &= \int_0^{\tau_h} P(\tau_{\text{cancer}}|k) P(n_p|\tau_{\text{cancer}}, r_p) d\tau_{\text{cancer}} \\ &\approx \int_0^{\tau_h} \frac{\tau_{\text{cancer}}^{k-1} k}{\tau_h^k} \frac{e^{-\tau_p \tau_p^{n_p}}}{n_p!} d\tau_{\text{cancer}} \\ &\approx 1/[*n_p!] \int_0^{\tau_h} e^{-\tau_p \tau_{\text{cancer}}^{k-1}} \tau_p^{n_p} d\tau_{\text{cancer}} \quad [\text{S24}] \\ &\approx k/[\tau_h^k n_p! r_p^k] \int_{\tau_p=0}^{\tau_p=\tau_h r_p} e^{-\tau_p \tau_p^{k-1+n_p}} d\tau_p \\ &\approx \binom{n_p+k-1}{n_p} k! \langle n_p^{\text{max}} \rangle^{-k} \gamma(k + n_p, \langle n_p \rangle^{\text{max}}) \end{aligned}$$

Where $\gamma(s, x)$ is the normalized incomplete gamma function defined previously (Eq. 8). In the above derivation, we eliminated a parameter by considering the quantity: $\langle n_p \rangle^{\text{max}} = \tau_h r_p$, which corresponds to the mean number of passengers expected for a person who lives until the maximum allowable time τ_h .

Lastly, it is important to remember that the total number of mutations $n_d + n_p$ is the expected number of passengers $P(n_p|k)$ plus the number of drivers k , which is constant. This is true for both predicted passenger distributions (Eqs. 24 and 23).

4 Materials & Methods

4.1 Simulations

All simulations were run using a previously-described first-order Gillespie algorithm [2]. Extinction was defined when a population declined to zero, while successful growth was defined when a population doubled. For each calculation of P_{cancer} , 3,000 simulation outcomes were averaged. To calculate predicted age-incidence curves, $2 \cdot 10^6$ simulations were used. All treatments began once a successful trajectory reached 10^6 cells and continued until extinction or until the population achieved $2 \cdot 10^6$ cells.

4.2 Inference of lesion formation rate

In the main text, we argue that r is at least $10 \text{ lesions} \cdot \text{year}^{-1}$ in breast epithelial. This lower-bound estimate was based on the assumption that $r = (10^3 \text{ breast epithelial stem cells per mouse [44]}) \times (2 \cdot 10^3 \text{ human breast epithelial stem cells per mouse breast epithelial stem cell}) \times (10^{-5} \text{ initiating mutations per cell per year [45]}) \approx 20$. Moreover, there are generally scores of lesions observed in normal breasts tissue [46, 47], which suggests that this estimate is close provided that lesions last for several years. This lead us to the conclusion that age-incidence curves cannot be explained by models which permit most lesions to eventually progress to cancer.

4.3 Analysis of cancer genomes

Driver genes were identified using MutSig [6] (for potential NSM drivers) and GISTIC 2.0 [48] (for SCNA drivers) in each dataset. To be defined as a ‘driver’, a mutation needed to arise in a gene with a Bonferroni-corrected enrichment p -value $\leq 5 \cdot 10^{-3}$. All other mutations were classified as ‘passengers’. These classifications were done in the articles where the SNMs and SCNAs were first identified in exome-sequenced tumor-normal pairs: breast [49], colon [16], lung [50], and skin [51] cancer (**Table S2**). MIN colorectal cancers were distinguished from non-MIN tumors in a previous study [16].

We chose to normalize observed and expected distributions (for both our model and the neutral passenger model) by their median because both models can adjust their median with a free parameter. In the neutral passenger model, this parameter is $\langle n_p^{\text{max}} \rangle$, while in our model, T_p/T_d and s_p could be simultaneously adjusted to fit the observed median without altering other properties of the distribution (see Eq. 1 in the main text).

There were also 6 breast cancers (5% of the dataset) with mutation totals greater than 4 times the median. All models in our study poorly explain these cancers, so we excluded them from our analysis. The breast cancer distribution was then compared to the various expected distributions using a violin plot. Error bar violin curves denoting the bottom 5th and top 95th percentile of the

observed distribution were generated by creating 10,000 violin plots from 10,000 bootstrapped resamplings of the observed distribution.

These analyses were repeated for the 11 cancer subtypes that currently have ≥ 100 sequenced exomes. Using a Kolmogorov-Smirnov test for quality-of-fit, we compared each observed distribution to our simulated distributions for various $s_d \in \{ 0.05, 0.1, 0.2, 0.4, 0.6, 0.8 \}$ and identified the best-fitting value of s_d . We then repeated these quality-of-fit tests for the driver-only model and identified the values of k that best explain the observed distributions (**Fig. S3, Table S4**). In general, our model explains the observed distributions more accurately than the driver-only model, with $s_d \approx 0.1-0.6$, while the driver-only model leads to unrealistically small values of k (between 1 and 2). In fact, for many cancer subtypes, our model explains the observed distribution to statistical resolution.

References

- [1] Gibson Ma, Bruck J (2000) Efficient Exact Stochastic Simulation of Chemical Systems with Many Species and Many Channels. *J Phys Chem A* 104:1876–1889.
- [2] McFarland CD, Korolev KS, Kryukov GV, Sunyaev SR, Mirny La (2013) Impact of deleterious passenger mutations on cancer progression. *Proc Natl Acad Sci USA* 110:2910–5.
- [3] Johnston MD, Edwards CM, Bodmer WF, Maini PK, Chapman SJ (2007) Mathematical modeling of cell population dynamics in the colonic crypt and in colorectal cancer. *Proc Natl Acad Sci USA* 104:4008–13.
- [4] Bota S, et al. (2001) Follow-up of bronchial precancerous lesions and carcinoma in situ using fluorescence endoscopy. *Am J Respir Crit Care Med* 164:1688–93.
- [5] Domingues JS (2012) Gompertz Model : Resolution and Analysis for Tumors. *J Math Model Appl* 1:70–77.
- [6] Lawrence MS, et al. (2013) Mutational heterogeneity in cancer and the search for new cancer-associated genes. *Nature* 499:214–8.
- [7] Shibata D, Peinado MA, Ionov Y, Malkhosyan S, Perucho M (1994) Genomic instability in repeated sequences is an early somatic event in colorectal tumorigenesis that persists after transformation. *Nat Genet* 6:273–81.
- [8] Nowak Ma, et al. (2002) The role of chromosomal instability in tumor initiation. *Proc Natl Acad Sci USA* 99:16226–31.
- [9] Haag-Liautard C, et al. (2007) Direct estimation of per nucleotide and genomic deleterious mutation rates in *Drosophila*. *Nature* 445:82–5.
- [10] Tong aH, et al. (2001) Systematic genetic analysis with ordered arrays of yeast deletion mutants. *Science* 294:2364–8.

- [11] Camps M, Herman A, Loh E, Loeb LA (2007) Genetic constraints on protein evolution. *Crit Rev Biochem Mol Biol* 42:313–26.
- [12] Geiler-Samerotte KA, et al. (2011) Misfolded proteins impose a dosage-dependent fitness cost and trigger a cytosolic unfolded protein response in yeast. *Proc Natl Acad Sci USA* 108:680–5.
- [13] Boyko AR, et al. (2008) Assessing the evolutionary impact of amino acid mutations in the human genome. *PLoS Genet* 4:e1000083.
- [14] Keightley PD, Kryukov GV, Sunyaev S, Halligan DL, Gaffney DJ (2005) Evolutionary constraints in conserved nongenic sequences of mammals. *Genome Res* 15:1373–8.
- [15] Kryukov GV, Schmidt S, Sunyaev S (2005) Small fitness effect of mutations in highly conserved non-coding regions. *Hum Mol Genet* 14:2221–9.
- [16] Cancer T, Atlas G (2012) Comprehensive molecular characterization of human colon and rectal cancer. *Nature* 487:330–7.
- [17] Futreal PA, et al. (2004) A census of human cancer genes. *Nat Rev Cancer* 4:177–83.
- [18] Beerenwinkel N, et al. (2007) Genetic progression and the waiting time to cancer. *PLoS Comput Biol* 3:e225.
- [19] Beckman RA, Loeb LA (2005) Negative clonal selection in tumor evolution. *Genetics* 171:2123–31.
- [20] S Datta R, Gutteridge A, Swanton C, Maley CC, Graham TA (2013) Modelling the evolution of genetic instability during tumour progression. *Evol Appl* 6:20–33.
- [21] Eisenberg E, Levanon EY (2013) Human housekeeping genes, revisited. *TIG* 29:569–74.
- [22] Vesely MD, Schreiber RD (2013) Cancer immunoediting: antigens, mechanisms, and implications to cancer immunotherapy. *Ann N Y Acad Sci* 1284:1–5.
- [23] Sheltzer JM, Amon A (2011) The aneuploidy paradox: costs and benefits of an incorrect karyotype. *Trends Genet* 27:446–53.
- [24] Bauer B, Siebert R, Traulsen A (2014) Cancer initiation with epistatic interactions between driver and passenger mutations. *J Theor Biol* pp 1–9.
- [25] Elgendy M, Sheridan C, Brumatti G, Martin SJ (2011) Oncogenic Ras-induced expression of Noxa and Beclin-1 promotes autophagic cell death and limits clonogenic survival. *Molecular cell* 42:23–35.

- [26] Michor F, Iwasa Y, Lengauer C, Nowak MA (2005) Dynamics of colorectal cancer. *Seminars in cancer biology* 15:484–93.
- [27] Vermeulen L, et al. (2013) Defining stem cell dynamics in models of intestinal tumor initiation. *Science* 342:995–8.
- [28] Kumar MS, et al. (2009) Dicer1 functions as a haploinsufficient tumor suppressor. *Genes Dev* 23:2700–2704.
- [29] Philipp-Staheli J, Payne SR, Kemp CJ (2001) p27ⁱ, supⁱ, kip1ⁱ/supⁱ: Regulation and function of a haploinsufficient tumor suppressor and its misregulation in cancer. *Exp Cell Res* 264:148–168.
- [30] Mao JH, et al. (2004) Fbxw7/cdc4 is a p53-dependent, haploinsufficient tumour suppressor gene. *Nature* 432:775–779.
- [31] Park BJ, et al. (2005) The haploinsufficient tumor suppressor p18 upregulates p53 via interactions with atm/atr. *Cell* 120:209–221.
- [32] Johnson T, Barton NH (2002) The effect of deleterious alleles on adaptation in asexual populations. *Genetics* 162:395–411.
- [33] Neher RA, Shraiman BI (2012) Fluctuations of fitness distributions and the rate of Muller’s ratchet. *Genetics* 191:1283–93.
- [34] Goyal S, et al. (2012) Dynamic mutation-selection balance as an evolutionary attractor. *Genetics* 191:1309–19.
- [35] Brunet E, Rouzine IM, Wilke CO (2008) The stochastic edge in adaptive evolution. *Genetics* 179:603–20.
- [36] Gordo I, Charlesworth B (2000) The degeneration of asexual haploid populations and the speed of Muller’s ratchet. *Genetics* 154:1379–87.
- [37] Nowak MA (2006) *Evolutionary dynamics: exploring the equations of life* (Belknap Press).
- [38] Good BH, Rouzine IM, Balick DJ, Hallatschek O, Desai MM (2012) Distribution of fixed beneficial mutations and the rate of adaptation in asexual populations. *Proc Natl Acad Sci USA* 109:4950–5.
- [39] Korolev KS, et al. (2012) Selective sweeps in growing microbial colonies. *Phys Biol* 9:026008.
- [40] Bachtrog D, Gordo I (2004) Adaptive evolution of asexual populations under Muller’s ratchet. *Evolution* 58:1403–13.
- [41] Hanahan D, Weinberg RA (2011) Hallmarks of cancer: the next generation. *Cell* 144:646–74.
- [42] Armitage P, Doll R (1954) The age distribution of cancer and a multi-stage theory of carcinogenesis. *Br J Cancer* 8:1–12.

- [43] Meza R, Jeon J, Moolgavkar SH, Luebeck EG (2008) Age-specific incidence of cancer: Phases, transitions, and biological implications. *Proc Natl Acad Sci USA* 105:16284–9.
- [44] Kordon EC, Smith GH (1998) An entire functional mammary gland may comprise the progeny from a single cell. *Development* 125:1921–30.
- [45] Lynch M (2010) Rate, molecular spectrum, and consequences of human mutation. *Proc Natl Acad Sci USA* 107:961–8.
- [46] Page DL, Dupont WD, Rogers LW, Rados MS (1985) Atypical hyperplastic lesions of the female breast. A long-term follow-up study. *Cancer* 55:2698–708.
- [47] Wellings SR, Jensen HM, Marcum RG (1975) An atlas of subgross pathology of the human breast with special reference to possible precancerous lesions. *J Natl Cancer Inst* 55:231–73.
- [48] Mermel CH, et al. (2011) GISTIC2.0 facilitates sensitive and confident localization of the targets of focal somatic copy-number alteration in human cancers. *Genome Biol* 12:R41.
- [49] Stephens PJ, et al. (2012) The landscape of cancer genes and mutational processes in breast cancer. *Nature* 486:400–4.
- [50] Ding L, et al. (2008) Somatic mutations affect key pathways in lung adenocarcinoma. *Nature* 455:1069–75.
- [51] Berger MF, et al. (2012) Melanoma genome sequencing reveals frequent PREX2 mutations. *Nature* 485:502–6.
- [52] Howlader N, et al. (2013) SEER Cancer Statistics Review 1975-2010.
- [53] Loeb La, Bielas JH, Beckman Ra (2008) Cancers exhibit a mutator phenotype: clinical implications. *Cancer Res* 68:3551–7; discussion 3557.
- [54] Sjöblom T, et al. (2006) The consensus coding sequences of human breast and colorectal cancers. *Science* 314:268–74.
- [55] Beroukhi R, et al. (2010) The landscape of somatic copy-number alteration across human cancers. *Nature* 463:899–905.
- [56] Cole AM, et al. (2010) Cyclin D2-cyclin-dependent kinase 4/6 is required for efficient proliferation and tumorigenesis following Apc loss. *Cancer Res* 70:8149–58.
- [57] Ciriello G, et al. (2013) Emerging landscape of oncogenic signatures across human cancers. *Nat Genet* 45:1127–1133.

5 Supplemental Figures & Tables

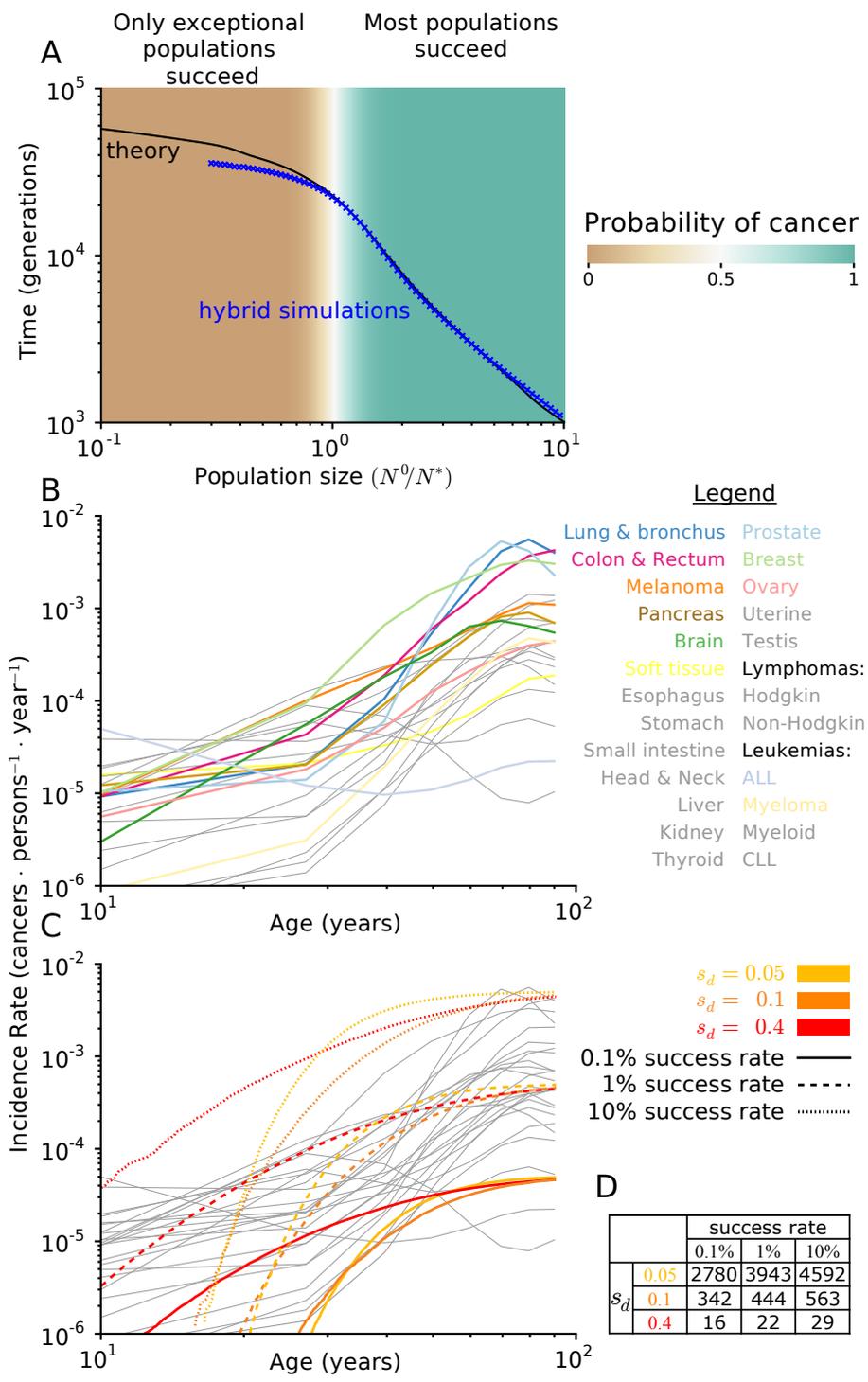


Figure S1 (*preceding page*): **Age-incidence curves in our model depend primarily on s_d and match observed age-incidence curves well at mid/late life.** **A** Mean waiting time to cancer $\overline{t_{\text{cancer}}}(x)$ decreases as initial population size $x = N^0/N^*$ increases. $\overline{t_{\text{cancer}}}(x)$ was solved from Eq. 5 using two methods: (1) stochastic ‘hybrid’ Gillespie simulations (Eq. 9), and (2) an analytical approximation (Eq. 10, labeled ‘theory’). Agreement between these two estimates suggests that our solution is accurate and a simplified analytical treatment of dynamics is possible (see **Estimating the mean time to progression** for details). These results demonstrate that the shape of our predicted age-incidence curves (below) should depend almost entirely on s_d and not x when $x < 1$, thereby simplifying interpretation of age-incidence curves. **B** Incidence rate versus age for the 25 most common cancers in the SEERs database [52]. Nearly all cancers show incidence rates that rise rapidly at mid-life, but then plateau at old-age. Leukemias have flatter curves, suggesting that they need fewer drivers for carcinogenesis. Only colorectal cancer does not plateau. Instead, it exhibits a power-law relationship for all ages. Some incidence curves flatten at young ages, which has been linked to germ-line predispositions to cancer that expedite progression. Neither our model nor the traditional neutral-model of attempt to explain these childhood occurrences. **C** The predicted age-incidence curves derived from simulations match observed age-incidence curves in most cancer subtypes when proper parameters are chosen. The slope of predicted age-incidence curves is described by s_d : a larger s_d causes the slope of age-incidence curves to decrease. The location in the plateau of age-incidence curves is described by the success rate of cancer progression P_∞ multiplied by the lesion formation rate r . These two parameters introduce, essentially, a Gauge freedom into our comparison of simulations with theory. Hence, it is most useful to simply think of only one additional parameter (after s_d): a overall height of the plateau. This effective Gauge freedom is evident by the fact that the incidence curves with the same s_d look approximately the same, irregardless of their success rate. Only their overall height changes (as predicted in **A**). For the predicted age-incidence curves plotted, $r = 5$. Of course this value, and the success rate, presumably vary considerably between cancer types. **D** The actual initial population size (N^0) needed to obtain various success rates of cancer progression from various s_d . Values were obtained by iteratively simulating various initial sizes until converging to an initial population size that led to the desired success rate. These values differed only mildly from the predictions of our analytical theory.

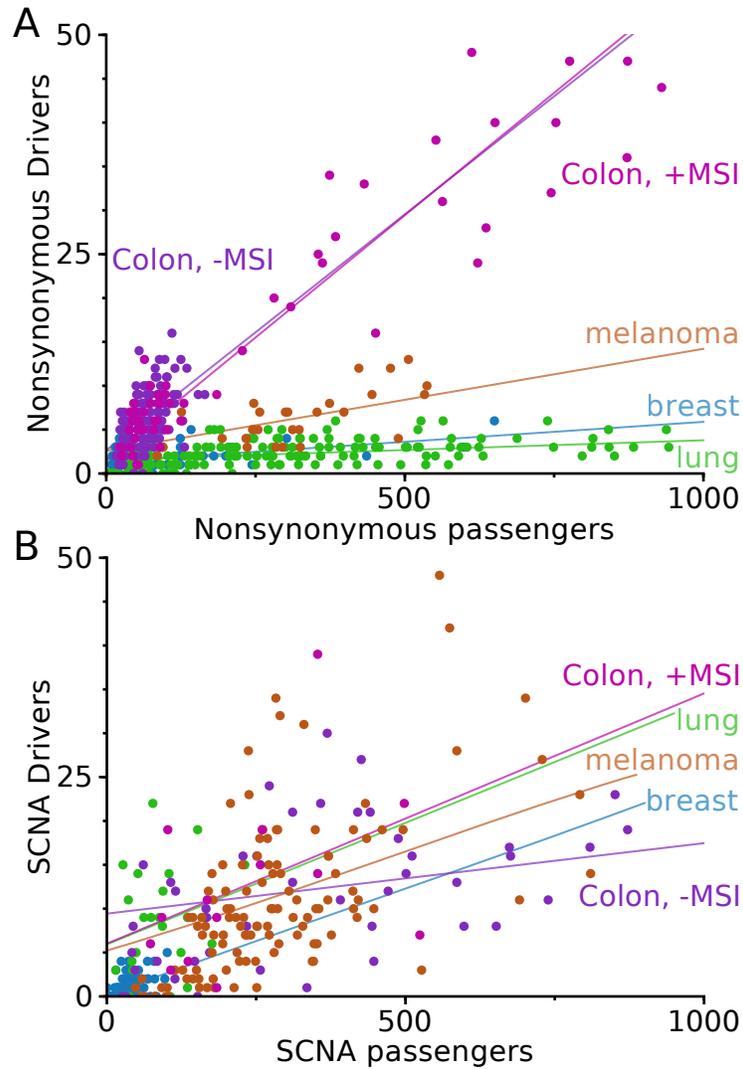


Figure S2: **Somatic Nonsynonymous Mutations (SNMs) and Somatic Copy Number Alterations (SCNAs) exhibit similar positive linear relationships among cancer subtypes.** **A** A positive linear relationship is observed between driver and passenger SNMs in all cancer subtypes studied here. This suggests that additional SNM passengers are being counterbalanced by additional drivers, and is consistent with our conclusions in the main text. Slope, y-intercepts, and the statistical significance of each best-fit line can be found in **Table S3**. **B** Positive linear relationship is also observed in SCNAs. The similar slopes and y-intercepts of SCNAs to SNMs supports our assumption that SCNAs and SNMs can be aggregated in analysis and modeling.

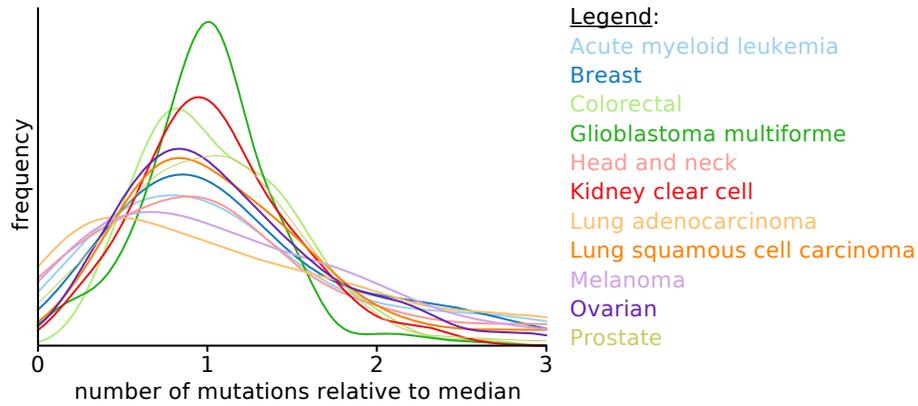


Figure S3: **Distribution of mutation totals in cancers is highly dispersed and positively skewed.** 11 cancer subtypes have 100 or more tumors sequenced via TCGA [6]. These subtypes all had widely-different distributions of mutation totals; however, they all appear to have a large degree of variance within their subtype and positive skew. Violin plots of each distribution are shown. In **Table S4**, we compared our model of cancer progression, for various s_d , alongside a traditional, neutral-passenger model of cancer progression (see **A traditional model of cancer progression with drivers and neutral passengers.**) to these distributions.

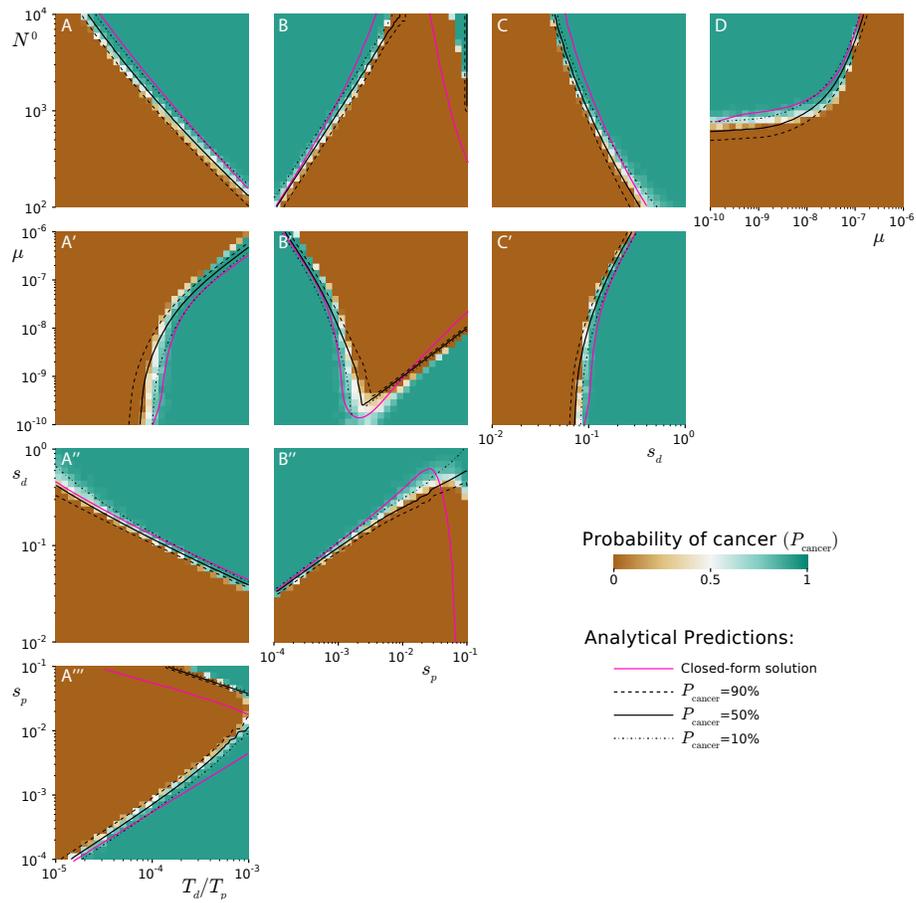


Figure S4 (*preceding page*): **Analytical framework predicts probability of cancer across parameter space.** The probability of progression, determined from the outcome of 3,000 simulations (for each data-square) propagated until extinction or rapid growth, across the parameter range of our model. In simulations, we observe parameters where progression occurs, fails, and is rare. A sophisticated analytical framework incorporating selection against passengers, hitchhiking of passengers onto driver mutations, and stochasticity in population size predicts observations well (black lines). This sophisticated analytical model uses two solutions for Muller’s Ratchet in various parameter regimes (see **Selection against passengers.**), an estimate of the quantity of hitchhiking passengers (see **Effects of passengers on driver fixation**), and a stochastic differential equation of the population size to estimate probabilities of progression (see **Estimating the probability of cancer**). A simplified framework, which offers a closed-form solution, is possible and works reasonably well (magenta). This solution differs from the more precise solution in two ways: (1) a novel, simplified estimate of Muller’s ratchet is used (Eq. 12), and (2) we neglect hitchhikers that accumulate after a new driver clone arises (i.e. $\delta p_S = 0$). **A-A'''** P_{cancer} increases for all parameters, as the relative target size of drivers T_d versus passengers T_p increases. **B-B'''** and **A'''** P_{cancer} exhibits a local minimum versus the selection against passengers s_p . When selection against passengers is very weak, passengers are effectively neutral. When selection against passengers is too strong, natural selection prevents passengers from accumulating. Deleterious passengers are most effective at preventing cancer when moderate in effect size. The local minimum suggests that there may be two types of cancers: those existing in an environment or genetic context where passengers are weak, perhaps buffered by an activated UPR; and those that succeed by exacerbating passengers’ deleterious effects, perhaps by decreasing their mutation rate (see **B'**). **C, C', A'', B''** The probability of cancer always increases with s_d . This parameter has a profound affect on cancer progression as it increases the benefit of drivers and their probability of fixation. Hence, the boundary between success and failure appears to be almost independent of the other parameters. **D** and **A'** An increasing mutation rate affects the probability of cancer very little at first; however, once it exceeds a critical value ($\mu^* = s_d/T_p$), the probability of cancer drops precipitously.

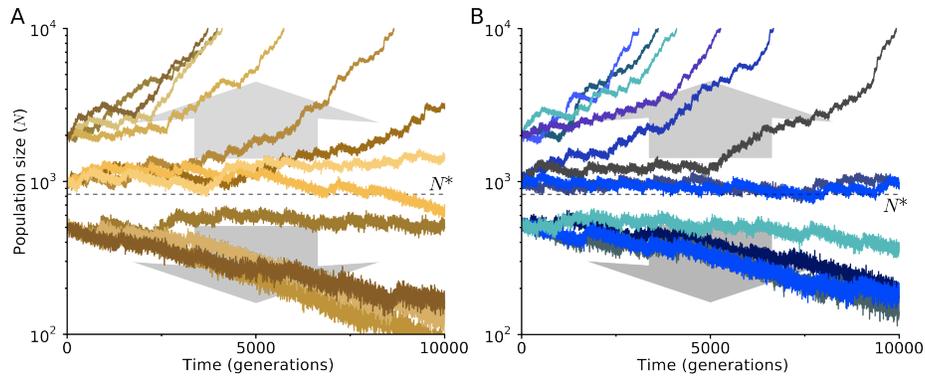


Figure S5: **Simulations exhibit path independence.** **A** The 12 trajectories from **Figure 1A**, initiated at $N^0 = \{500, 1000, 2000\}$. **B** An additional 12 trajectories, initiated at various N^0 , but plotted once they cross $N = \{500, 1000, 2000\}$. Populations that crossed $N = 500$ and $N = 2000$ were initiated at $N^0 = 1000$, while populations that crossed $N = 1000$ were initiated at $N^0 = 500$. Dynamics in **A** and **B** appear identical, demonstrating that populations beginning at different initial sizes N^0 will behave similarly, if they have the same current size. Thus, populations exhibit path independence and can be fully described by one state variable: $x = N/N^0$.

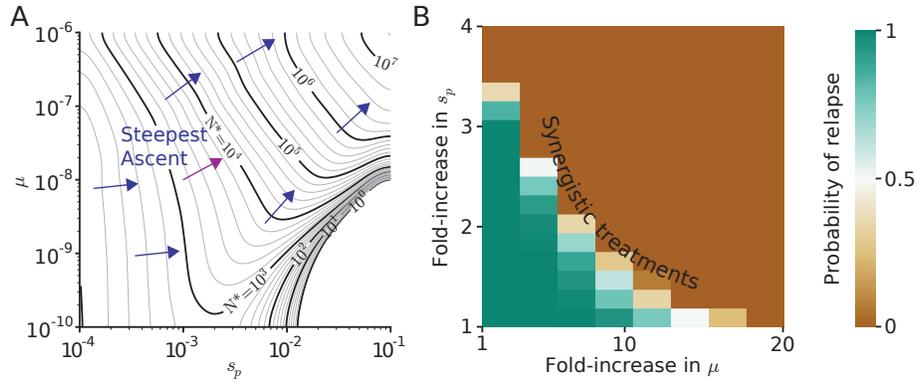


Figure S6: Combination treatments that increase mutation rate and selection against passengers work best. **A** Using the analytical theory describe in **An analytical model of dynamics** and **Fig. S4**, we plotted the critical population size N^* across evolutionary parameters as a contour plot. Optimal therapy, from an evolutionary perspective, should increase N^* along its gradient of steepest ascent (blue lines). From this 3-Dimensional perspective the interplay between μ and s_p is evident. For cancers with low mutation rates, only weak passengers (low s_p) can fixate. Thus, these cancers should be more immune to drugs that increase s_p . Cancers with high mutation rates fixate all passengers, making passenger-targeted therapies highly effective. At intermediate mutation rates, the most effective treatment would moderately increase both the mutation rate and s_p . **B** Via simulations, we tested our prediction that the gradient of steepest ascent is optimal for the magenta-colored vector in **A**. 50 cancers with $\mu = 10^{-8}$, $s_p = 0.001$ grown to 10^6 cells were treated with combinations of mutagenic and s_p increasing therapy. Indeed, moderate increases in both parameters were more effective than would be expected from the lone treatments, thus confirming our prediction. These results underscore the importance of combinatorial therapies and evolutionary modeling for cancer treatment.

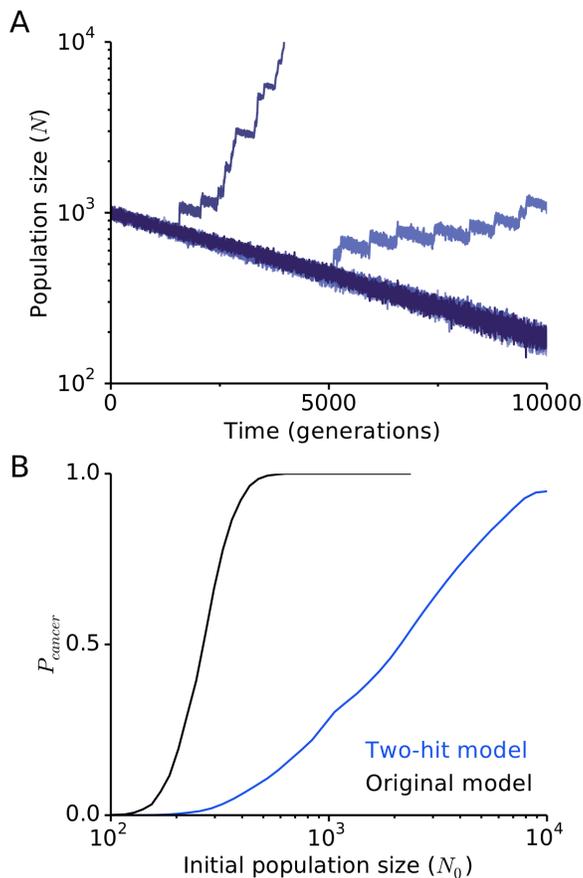


Figure S7: **A two-hit driver model with deleterious passengers also experiences a population size-dependent barrier to progression.** **A** We extended our original model to consider cancers where the first driver mutation confers no fitness benefit, the second driver confers a fitness benefit of $2s_d$, and all remaining drivers confer a fitness benefit of s_d . All other properties of the model, including deleterious passengers and a dynamic population size, remain intact (see **A two-hit model of cancer progression** for details). For this figure $s_d = 0.2$, $s_p = 0.001$, $\mu_d = 1.4 \times 10^{-5}$, and $\mu_p = 0.1$. Trajectories exhibit a period of stasis where the first driver accumulates (according to neutral dynamics) before the second driver arrives. After this period of stasis, most trajectories progress to cancer. Passengers continue to accumulate and cause extinction in this model. **B** Like the original model, we observed a population size-dependent barrier to progression in the two-hit model. However, because of the long and variable period of stasis where passengers continue to accumulate, the critical population size $N^{*,two-hit}$ is larger and the transition from the extinction and growth regimes is slower. This is more-or-less equivalent to our model with a larger value of s_d and smaller value of T_d . See **A two-hit model of cancer progression** for a explanation of these findings using our analytical model of progression.

Table S1: **Evolutionary parameters explored in this study.**

Parameter	Symbol	Estimate	Range	Citation
Mutation rate	μ	10^{-8}	10^{-10} - 10^{-7}	[53]
Driver loci	T_d	700	70 - 7,000	[17, 18, 54]
Passenger loci	T_p	$5 \cdot 10^6$	$5 \cdot 10^5$ - $5 \cdot 10^7$	[13, 19]
Driver strength	s_d	0.1	0.001 - 1	[18, 55]
Passenger strength	s_p	0.001	10^{-4} - 10^{-1}	[12]
Initial population size	N^0	1000*	100 - 10,000	[56]

*Estimated from labeled populations in mice colonic crypts 2 weeks after an initiating *APC* deletion was induced.

We explored our evolutionary model incorporating driver and passenger mutations across a broad range of parameters. The ranges were motivated by literature estimates discussed previously [2] and in **Parameter range investigated**. Note that in simulations $\mu_d = \mu T_d$ and $\mu_p = \mu T_p$, hence the entire phase space can be explored by only altering μ and T_d/T_p , as altering all three parameters is redundant. In **Figure 2** we compare our model to epidemiological and genomic data and affirm that these prior published estimates explain the new data well.

Table S2: **Average number of driver and passenger mutations by tumor type.**

Cancer	NSM		SCNA	
	Drivers	Passengers	Drivers	Passengers
breast	1.7	70.8	1.0	34.6
lung	2.3	348.6	8.4	89.5
colon, MIN ⁻	8.8	114.0	14.1	583.5
colon, MIN ⁺	28.8	489.0	12.7	235.1
melanoma	7.0	379.6	12.6	324.7
all	9.1	272.8	8.8	258.9
Max	28.8	489.0	14.1	583.5
Min	1.7	70.8	1.0	34.6

The total number of identified Somatic Nonsynonymous Mutations (SNMs) and Somatic Copy Number Alterations (SCNAs) for various tumor-normal paired sequences from various tissues of origin: 100 breast [49], 183 lung [50], 159 Colon without Micro-satellite INstability (MIN⁻), 64 Colon with Micro-Satellite Instability (MIN⁺) [16], and 121 melanomas [51].

Table S3: Linear relationship between drivers and passengers cannot be explained by other tumor properties.

Cancer	Pearson's r	p -value*	N †	Spearman's ρ	slope‡	y-intercept
Drivers versus Passengers						
breast	0.423	$< 10^{-4}$	100	0.413	0.006	2.02
lung	0.368	0.08	24	0.998	0.005	8.63
colon, MIN ⁻	0.624	$< 10^{-4}$	49	0.985	0.009	17.11
colon, MIN ⁺	0.916	$< 10^{-5}$	14	0.999	0.047	6.50
melanoma	0.749	$< 10^{-5}$	29	0.995	0.015	3.69
All	0.937	$< 10^{-99}$	217	0.992	0.042	-3.81
SNM drivers versus SNM passengers						
breast	0.390	$< 10^{-4}$	100	0.178	0.005	1.34
lung	0.587	$< 10^{-17}$	183	0.579	0.002	1.56
colon, MIN ⁻	0.990	$< 10^{-134}$	159	0.569	0.054	2.65
colon, MIN ⁺	0.994	$< 10^{-60}$	64	0.918	0.056	1.56
melanoma	0.878	$< 10^{-9}$	29	0.974	0.012	2.59
All	0.924	$< 10^{-223}$	536	0.592	0.050	-4.45
SCNA drivers versus SCNA passengers						
breast	0.443	$< 10^{-5}$	100	0.433	0.024	0.17
lung	0.253	0.23	24	0.998	0.028	5.94
colon, MIN ⁻	0.770	$< 10^{-9}$	49	0.984	0.008	9.44
colon, MIN ⁺	0.424	0.13	14	0.994	0.029	6.01
melanoma	0.559	$< 10^{-10}$	121	0.663	0.023	5.23
All	0.573	$< 10^{-27}$	309	0.962	0.012	5.76
SNMs versus SCNAs						
breast	0.052	0.61	100	0.149	0.237	64
lung	0.169	0.43	24	-0.548	1.268	334
colon, MIN ⁻	-0.080	0.58	49	-0.068	-0.021	137
colon, MIN ⁺	-0.265	0.36	14	0.045	-0.981	838
melanoma	-0.114	0.56	29	0.176	-0.183	431
All	0.331	$< 10^{-6}$	217	-0.089	0.631	133
Drivers versus Pathological Grade						
breast	0.163	0.10	100	0.113	0.067	2.25
lung	-0.048	0.83	22	0.024	0.006	2.13
colon, MIN ⁻	-0.187	0.20	48	0.072	0.012	2.67
colon, MIN ⁺	-0.121	0.68	14	-0.338	0.004	3.09
melanoma	0.221	0.35	20	0.120	0.025	1.83
All	0.018	0.80	204	0.054	0.001	2.37
SNMs versus Pathological Grade						
breast	0.217	0.03	100	0.444	0.001	2.33
lung	0.193	0.02	158	0.235	0.000	1.79
colon, MIN ⁻	-0.084	0.30	156	0.039	0.000	2.51
colon, MIN ⁺	-0.045	0.73	63	-0.023	0.000	2.50
melanoma	0.119	0.62	20	0.114	0.000	2.06
All	-0.012	0.80	497	0.147	0.000	2.28
SCNAs versus Pathological Grade						
breast	0.248	0.01	100	0.235	0.007	2.18
lung	0.170	0.45	22	0.026	0.002	1.85
colon, MIN ⁻	-0.166	0.26	48	0.081	0.000	2.48
colon, MIN ⁺	-0.054	0.85	14	-0.333	0.000	2.99
melanoma	0.109	0.31	88	-0.253	0.000	2.06
All	-0.067	0.27	272	-0.112	0.000	2.38

Negative values are in gray.

*Statistically significant ($p < 0.05$) relationship are in bold

† Number of samples compared

‡ Denotes s_p/s_d when comparing drivers to passengers

Table S3 (*preceding page*): We observe a linear relationship between drivers and passengers, predicted by our model. Above the thick black line are relationships that robustly covary, while the bottom half contains relationships that are generally insignificant. In our model, driver’s and passenger’s linear relationship results from their competing effect (additional deleterious passengers must be overcome by additional drivers), however alternate factors might have explained this relationship. In particular, we were concerned that the mutation type, mutation rate, or aggressiveness of the tumor could also explain the observed relationship. Thus, we used the analyses above to test these possibilities and found that the data does not support these competing hypotheses. Our rationale for the other competing hypotheses and why we reject them: (a) SCNAs and SNMs might have drastically different effects on cancer progression and undermine our model. The slope and y-intercept between drivers and passengers is *approximately* equal in SCNAs and SNMs, suggesting the relative fitness effects of these mutations is similar. (b) Some cancers might progress via CIN, while others progress via an elevated point mutation rate. If so, a negative correlation between SCNAs and SNMs within tumor subtypes would be expected, which has been observed previously in a pan-cancer study [57] and within the aggregate colorectal dataset. However, this does not appear to be so in other tumor types, nor in colorectal cancer after segregation according to MIN phenotype. Thus, the observed patterns are not explicable by varying mutational mechanisms. (c) The relationship between drivers and passengers might be a result of variation in mutation rate. Variation in the mutation rate should only alter the waiting time to cancer in the neutral-passenger model, and not alter mutation totals. Nevertheless, if variation in the mutation rate could explain the correlation between drivers and passengers, then stratifying tumors by their mutation rate should reduce the correlation. Because the relationship between drivers and passengers is persistent and strong within the MIN⁺ and MIN⁻ subtypes—expected to have and not-have a mutator phenotype—we reject this hypothesis. (d) Tumors with more drivers and passengers might simply be more evolutionarily advanced. Suppose some cancers are detected and sequenced later than others. These late cancers would not only possess additional drivers, but also additional passengers, even if passengers were neutral; thus, retaining the correlation between drivers and passengers. However, late-detected tumors with additional drivers should also be more advanced and more aggressive. We find that a tumor’s pathological grade is uncorrelated with the number of drivers, refuting this possibility. Pathological grade was quantified by converting roman numerals into a linear scale (i.e. A Stage IV tumor corresponds to an aggressiveness of 4). Many tumors had intermediate grades that were given corresponding fractional values (e.g. a Stage IIIa tumor was translated into a 3.0, a Stage IIIb was given 3.3, and a Stage IIIc was given 3.7). Because this quantification of tumor grade may distort the scale of aggressiveness, Spearman’s Rank correlations are provided. For completeness, we have also show the relationship between SNMs and Pathological Grade and SCNAs and Pathological Grade.

Table S4: **Kolmogorov-Smirnov goodness of fit estimates of s_d (our model) and k (traditional model) for sequenced cancer subtypes.**

Tissue	N	s_d	Our model p -value*	k	Traditional model p -value*
Acute myeloid leukemia	132	0.8	0.802	1	0.072
Breast	120	0.6	0.860	1	0.045
Colorectal	230	0.2	0.956	2	0.011
Glioblastoma multiforme	219	0.1	0.070	3	0.187
Head and neck	178	0.8	0.120	1	0.077
Kidney clear cell	214	0.2	0.485	2	0.066
Lung adenocarcinoma	333	0.8	3.71×10^{-5}	1	4.35×10^{-7}
Lung squamous cell	178	0.4	0.752	1	1.10×10^{-3}
Melanoma	121	0.8	0.267	1	0.073
Ovarian	385	0.4	0.833	1	4.54×10^{-6}
Prostate	221	0.4	0.134	1	0.018

*Two-sided p-value that the observed and expected distributions are identical

11 cancer subtypes have 100 or more tumors sequenced via TCGA [6]. These subtypes all had widely-different and broad distributions of mutation totals (**Fig. S3**). We then compared our model of cancer progression for various $s_d \in \{ 0.05, 0.1, 0.2, 0.4, 0.6, 0.8 \}$ and **A traditional model of cancer progression with drivers and neutral passengers** to these distributions. For comparison, we divided all predicted and observed mutation totals by their median and discarded outliers ≥ 4 times the median, as various confounding factors affect the distributions by a multiplicative constant (see **Supplemental Methods**). We then selected our best fitting model, and the best fitting traditional model (largest D statistic) using a Kolmogorov-Smirnov goodness of fit test and displayed the best fitting estimates of s_d or k alongside their quality of fit. From **Fig. S3**, note that Acute myeloid leukemia and Glioblastoma multiforme have *extremely* dispersed distributions that are unlike the other subtypes. We note several observations from this data: (1) our model explains most cancer subtypes well and outperforms the traditional model for all subtypes examined, excluding Glioblastoma multiforme; (2) $s_d \approx 0.2 - 0.6$ for most subtypes examined; (3) the best fitting traditional models assume that an unrealistically-low number of drivers are needed for progression ($\sim 1 - 2$).